WORKSHOP GARR 2020

GarrLab: giochiamo… ‘sul serio’?

(aka: il SIEM in cloud [GARR])

Damiano Verzulli
damiano@verzulli.it

APM GARR
Università “G. d'Annunzio” di Chieti-Pescara
http://www.unich.it

APM GARR
ICRANET Research Center
Pescara
http://www.icranet.org

Perchè questo intervento?

- nel termine **SIEM** la “**S**” (Security) non può esistere senza la “**I**” (Information) e la “**E**” (Event).
- “**I**” ed “**E**”, inoltre, vivono di vita propria;
- è soltanto aggiungendo la “**M**” (Management) che ingegnerizziamo la gestione di “**I**” ed “**E**” e quindi, finalmente, poniamo le basi per un’analisi (di **I** ed **E**) in vari ambiti, inclusa la **S**icurezza

L’oggetto della presentazione è un “attrezzo” (log-biter) che implementa la “**M**” di “**I**” e di ”**E**”

Perchè qui? Perchè ora?

“…il SIEM è una enorme lente d'ingrandimento: come ogni sistema di analisi vi fa vedere cose che voi umani non volete vedere …” - Simone Bonetti – CERT-UniBo - 23/04/2020 20:38

- la **mia sensazione** è che in molti – **specie negli Atenei non-grandi** – comincino a sentire l’esigenza di avere questa “lente”;
- Molti (dei “molti” di cui sopra) faticano a percepire la reale “potenza” di questa lente. Viceversa, intuiscono facilmente che:
  - la lente “va messa a fuoco” (altrimenti non serve a nulla)
  - la lente è pesante e difficile da pulire (HW, SW e storage, rispetto allo status-quo);
  - le lenti commerciali sono fuori dalla portata del proprio budget

Risultato: nella quasi totalità degli Atenei medio-piccoli pur essendoci “I” ed ”E”, manca la “M” e, quindi, non può esserci un SIEM. È un problema da WS GARR :-)

Obiettivo?

- L'obiettivo è quello di rimuovere (o almeno minimizzare) quei fattori che impediscono la diffusione di **M** negli Atenei medio-piccoli
- Due macro-tipologie di ostacoli:
  - **1 - Problemi "Tecnici"**: i numeri ci dicono che anche negli Atenei piccoli, NON è possibile approcciare "**M**(**I**,**E**)" utilizzando tecnologie classiche (DBMS).

Servono "nuove tecnologie" (big-data) accompagnate da una parziale riorganizzazione dei sistemi ("**I**" ed "**E**" vanno identificati e raccolti)

NON è realistico ipotizzare che tutto ciò accada in periferia, di moto proprio

Obiettivo?

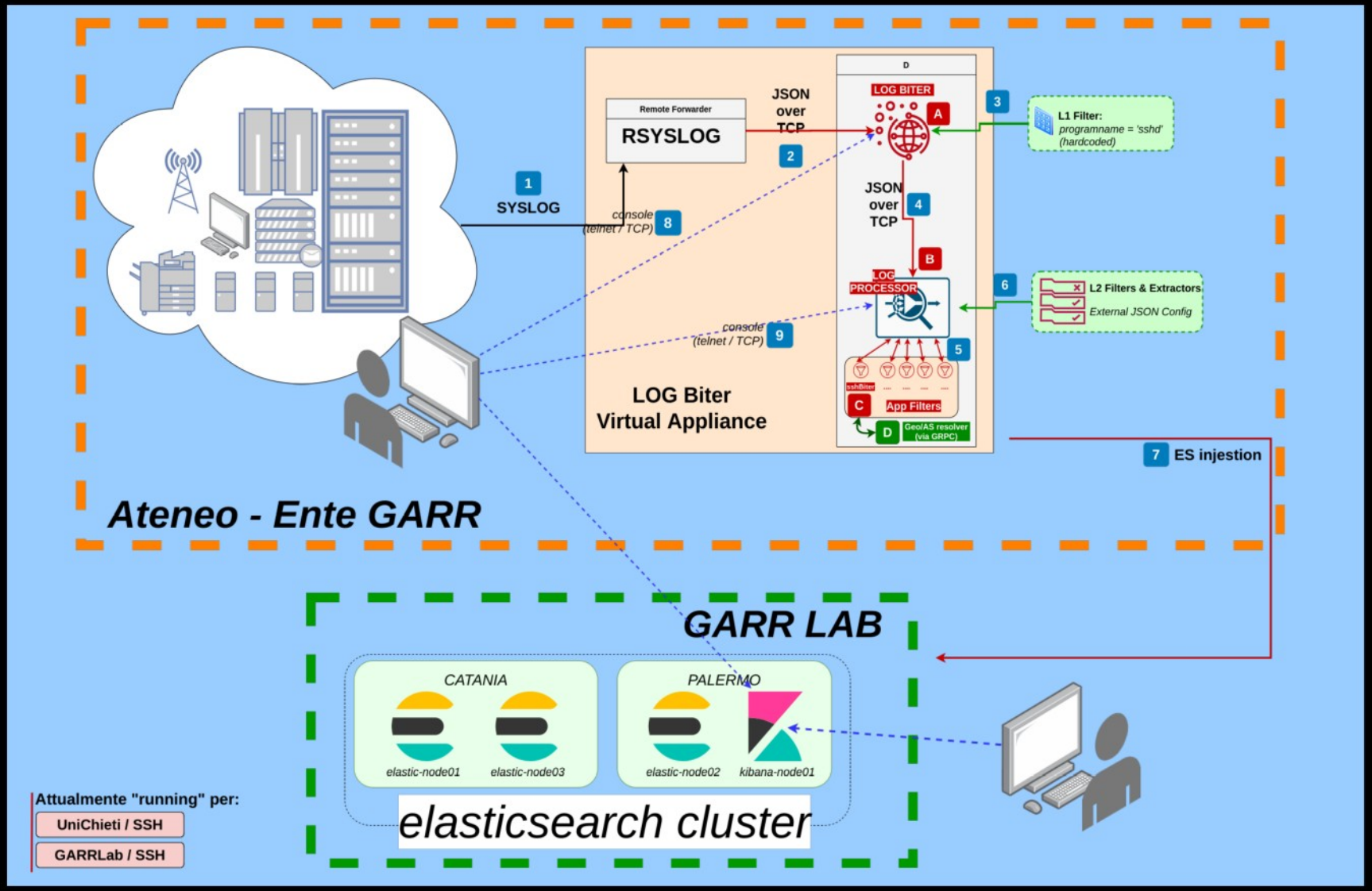
- (...continua...)
- **2 - Problemi "giuridico/organizzativi"**: all'interno di "E", spesso, ci sono diverse "I".
  Alcune di queste "I" sono tutelate dalla norma (GDPR) e vanno "trattate" in modo specifico (pseudononimizzazione; crittografia; deleghe al trattamento; modalità e tempi di conservazione; etc.)

Probabilmente questi fattori stanno determinando un certo "ritardo" della proliferazione di soluzioni "in outsourcing"

E GARRLab?

- In ambito GarrLab si sono spese un po’ di energie per allestire un POC che eliminasse i problemi “tecnici” lato Atenei, ossia:
  - ospitare “M” in GarrLab (ossia, su Cloud GARR), quindi niente carico aggiuntivo in Ateneo;
  - trasportare “E” ed “I”, dagli Atenei al GarrLab in modo sicuro (100% trasporto su infrastruttura GARR - “trusted by design”)
  - pre-processare “I” direttamente in Ateneo, per aggiungere i livelli di “pseudononimizzazione”/”anonimizzazione”/ ”crittografia” desiderati dall’Ateneo

Dettagli, please....



Dettagli, please....

```
biter@GARRLab> show counters
Log monitor counters:
    12390958: in_total
           0: in_filtered
           0: in_processed
       45516: in_sent
    12345442: in_skipped
           0: in_JSON_parse_errors
           0: processorMissed

Log monitor - Module counters:
             45516: sshBiter

Started on [Fri Oct 23 2020 12:08:16 GMT+0200
(Central European Summer Time)]
biter@GARRLab>
```

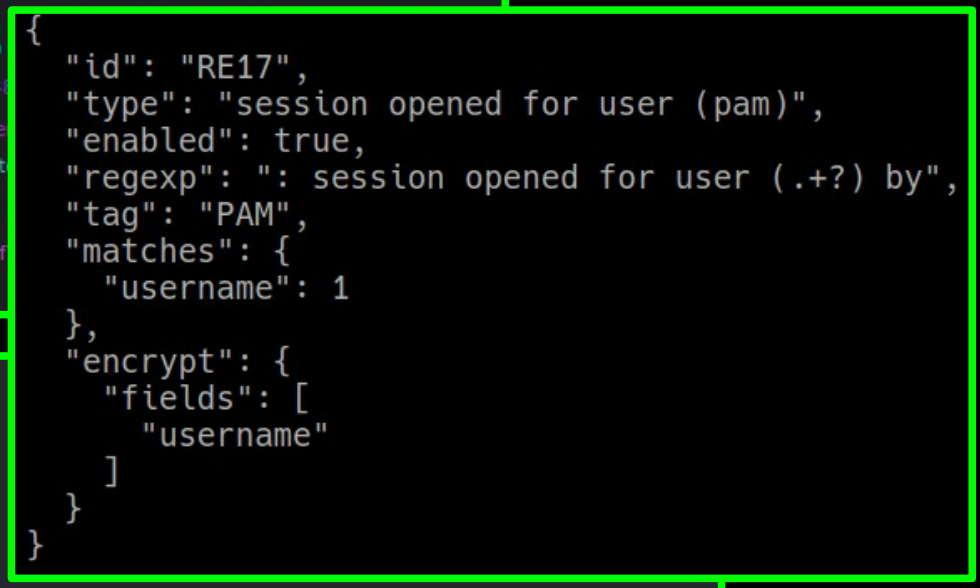
```
processor@GARRLab>

processor@GARRLab>

processor@GARRLab>

processor@GARRLab> show counters
           0: in_filtered
           0: in_processed
           0: in_sent
           0: in_JSON_parse_errors
       45520: in_JSON_parse_OK
Module counters:
             45520: sshBiter

processor@GARRLab>
```

```
biter@UNICH> show counters
Log monitor counters:
   904144877: in_total
           0: in_filtered
           0: in_processed
     8491450: in_sent
   895653426: in_skipped
           1: in_JSON_parse_errors
           0: processorMissed

Log monitor - Module counters:
             5508289: dhcpBiter
             2983161: sshBiter

Started on [Sun Oct 11 2020 07:36:29 GMT+0000 (Co
ordinated Universal Time)]
biter@UNICH>
```

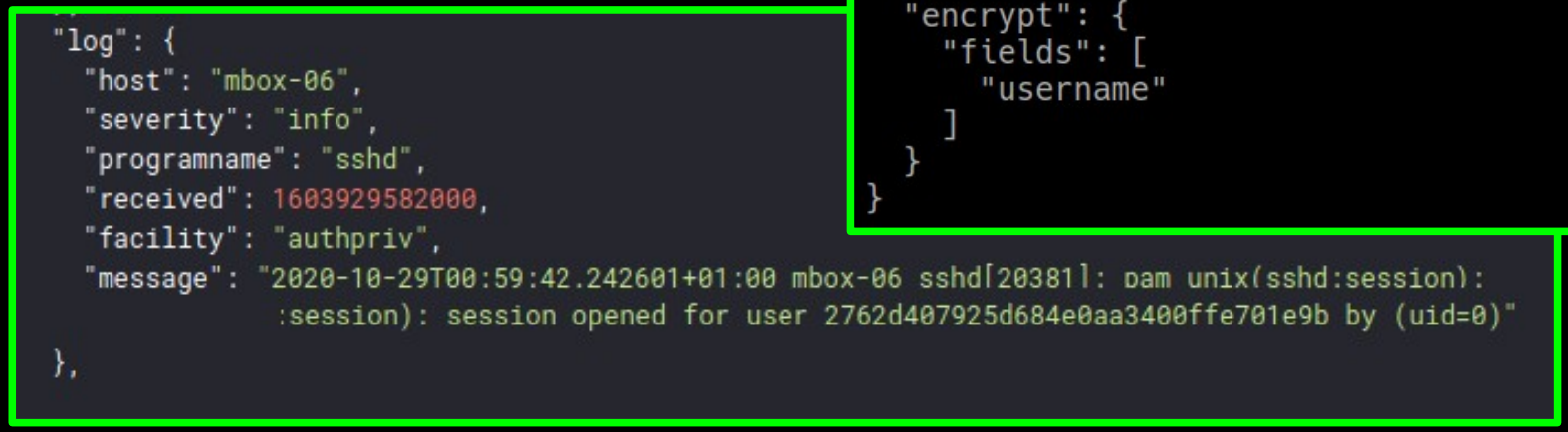
```
***          Log-Biter - PROCESSOR
***
***     v. 2020-10-11_01 - Console Server
***
**********************************************
***

processor@UNICH> show counters
           0: in_filtered
           0: in_processed
           0: in_sent
           0: in_JSON_parse_errors
     8491408: in_JSON_parse_OK
Module counters:
             5508261: dhcpBiter
             2983147: sshBiter

processor@UNICH>
```

aka

Dettagli, please.... (con la S)



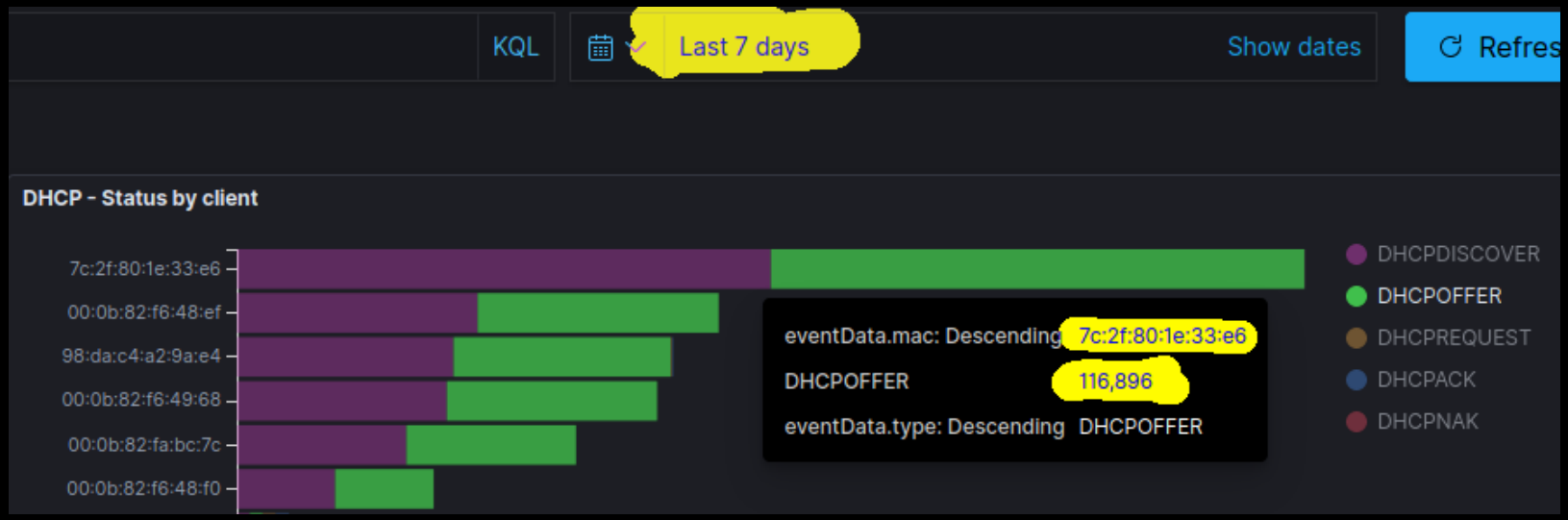
```
{
  "id": "RE17",
  "type": "session opened for user (pam)",
  "enabled": true,
  "regexp": ": session opened for user (.+?) by",
  "tag": "PAM",
  "matches": {
    "username": 1
  },
  "encrypt": {
    "fields": [
      "username"
    ]
  }
}
```



```
"log": {
  "host": "mbox-06",
  "severity": "info",
  "programname": "sshd",
  "received": 1603929582000,
  "facility": "authpriv",
  "message": "2020-10-29T00:59:42.242601+01:00 mbox-06 sshd[20381]: pam_unix(sshd:session):
              :session): session opened for user 2762d407925d684e0aa3400ffe701e9b by (uid=0)"
},
```

```
function encrypt(aText, aCipherKey, aInitVector) {
  let cipher = crypto.createCipheriv('aes-256-cbc', Buffer.from(aCipherKey), aInitVector);
  let encrypted = cipher.update(aText, 'utf8', 'hex');
  encrypted += cipher.final('hex');
  return encrypted;
}
```

Dettagli, please.... (senza la S)




```
  },
  {
    "id": "DH9",
    "type": "DHCPOFFER",
    "enabled": true,
    "sample": "Oct 10 22:49:00 RELA
GE= DHCPOFFER on 172.16.3.204 to 0
    "regexp": " DHCPOFFER on (\\d+
    "matches": {
      "ip": 1,
      "mac": 2,
      "eth": 3
    }
  },
  {
```

Per cifrare il MAC sarebbe sufficiente aggiungere la sezione ”encrypt/fields/mac”

Quindi è tutto risolto?

- **Assolutamente NO!** La componente “giuridico/organizzativa” è **fondamentale** e, a differenza dell’altra (quella “tecnica”) non abbiamo la capacità di affrontarla adeguatamente. Ad esempio:

- posso mandare i LOG dei miei sistemi e delle mie applicazioni a GARR? Se si…
- posso mandare gli “username” dei miei utenti, in chiaro, su quei log? No? E se li cifro con una chiave a disposizione solo del Responsabile IT “interno”?
- Posso mandare gli IP? Tutti? Solo gli RFC1918?
- Posso mandare i LOG della centrale telefonica (VoIP)?
- Posso mandare gli eventi generati dal file-server dell’amministrazione?
- etc. etc.

Please, help us!

L’idea di centralizzare sull’infrastruttura GarrLab il ricettore M(I,E) può avere un senso reale solo se è definita con:

- i livelli “di indirizzo IT” degli Atenei (RTD, DPO, CISO, etc.)
- e se è condivisa dalla governance (Rettore; Direttore Generale).

Senza l’endorsement di questi soggetti:

- ogni esercizio che nasce dal basso non può che restare un “esercizio” tecnico;
- la diffusione di servizi “SIEM-as-a-service” all’interno dei nostri Enti è solo questione di tempo (...servirà soltanto qualche “incidente serio”).

Grazie!

Parliamone!